Improving peptide fragmentation by N-terminal derivatization with high proton affinity.

AUTHOR(S):

Miyashita, Masahiro; Hanai, Yosuke; Awane, Hiroyuki; Yoshikawa, Toru; Miyagawa, Hisashi

CITATION:

Miyashita, Masahiro ...[et al]. Improving peptide fragmentation by N-terminal derivatization with high proton affinity.. Rapid communications in mass spectrometry 2011, 25(9): 1130-1140

ISSUE DATE:

2011-05-15

URL:

http://hdl.handle.net/2433/202058

Improving peptide fragmentation by N-terminal derivatization with high proton affinity

Short title: Peptide sequencing by derivatization with high proton affinity

Masahiro Miyashita*, Yosuke Hanai, Hiroyuki Awane, Toru Yoshikawa and Hisashi Miyagawa

Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

E-mail: miyamasa@kais.kyoto-u.ac.jp

Abstract. An improved method of *de novo* peptide sequencing based on mass spectrometry using novel N-terminal derivatization reagents with high proton affinity has been developed. The introduction of a positively charged group into the N-terminal amino group of a peptide is known to enhance the relative intensity of b-ions in product ion spectra, allowing the easy interpretation of the spectra. However, the physicochemical properties of charge derivatization reagents required for efficient fragmentation remain unclear. In this study, we prepared several derivatization reagents with high proton affinity, which are thought to be appropriate for peptide fragmentation under low-energy CID conditions, and examined their usefulness in *de novo* peptide sequencing. Comparison of the effects on fragmentation among three derivatization reagents having a guanidino or an amidino moiety, which differ in proton affinity, clearly indicated that there was an optimal proton affinity for efficient fragmentation of peptides.

Among reagents tested in this study, derivatization with 4-amidinobenzoic acid brought about the most effective fragmentation. This derivatization approach will offer a novel *de novo* peptide sequencing method under low-energy CID conditions.

Recent advances in various tandem mass spectrometric techniques have enabled the facile determination of the amino acid sequences of peptides[1-3]. In addition, the development of comprehensive databases of protein primary sequences predicted from the genomic sequences of various organisms has boosted the efficiency with which proteins can be identified by mass spectrometry (MS)[4]. With this database-dependent method, the assignment of each fragment ion in a product ion spectrum is not necessary, and identification is basically achieved by comparing the observed product ion spectra of enzymatically-digested peptides with those obtained by theoretical calculations based on primary sequences in databases using the Mascot[5], SEQUEST[6] or OMSSA[7] search engine. As primary sequence databases continue to grow, database-dependent methods are expected to provide more powerful and convenient tools for analyses of peptides/proteins; however at present, their coverage is limited to the components of major organisms in most cases. Amino acid sequencing without the assistance of a database (*de novo* sequencing) is still essential in the identification of peptides and proteins from unexplored organisms.

De novo sequencing can be carried out by interpreting mass differences between a series of consecutive ions of N- or C-terminal fragments generated by collision-induced dissociation (CID)[8]. Since tryptic peptides always contain Arg or Lys residues at the C-terminus, y-ions are preferentially generated, allowing the easy interpretation of spectra. However, when proteins are digested with other enzymes such as chymotrypsin, basic residues are located at various positions in digested peptides. Fragmentation of these peptides gives complicated and often incomplete product ion spectra, making *de*

novo sequencing difficult. This is also true for peptidomic analysis, in which peptides are generally analyzed without trypsin digestion[9]. Newly developed mass spectrometric techniques for better fragmentation, such as electron capture dissociation (ECD)[10] or electron transfer dissociation (ETD)[11], and for highly accurate mass measurements, such as Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry[12] or Orbitrap[13], have greatly improved the efficiency of *de novo* sequencing; however, there still remain difficulties in obtaining mass spectra that allow easy identification of key fragment ions for sequencing. The difficulties come mainly from incomplete fragmentation of the peptide backbone.

To overcome these problems, chemical modifications at the N-terminus of peptides have been attempted for more regular and “informative” fragmentation[14]. The introduction of a positively charged group, such as quaternary or tertiary ammonium and quaternary phosphonium groups, at the N-terminus is known to enhance the relative intensity of N-terminal fragment ions (b-ions) in product ion spectra, significantly facilitating the formation of whole series of b-ions[15-18]. Similarly, it has been reported that the introduction of a negatively charged group at the N-terminus enhances the generation of C-terminal fragment ions (y-ions) with suppression of the occurrence of b-ions[19]. This modification is particularly effective for tryptic peptides, in which y-ions are predominantly generated.

Tris(2,4,6-trimethoxyphenyl)phosphonium acetic acid (TMPP-ac), which contains a quaternary phosphonium group, has been widely used for the derivatization to provide an N-terminal positive charge[20]. This modification favors the generation of N-terminal

fragment ions under high-energy CID conditions by virtue of a fixed positive charge at the N-terminus; however, derivatization with a fixed charge has been shown to reduce the efficiency of fragmentation under low-energy CID conditions, which are commonly used in tandem quadrupole or quadrupole ion-trap mass spectrometers, and to often induce an almost exclusive cleavage at the amide bond of Asp on its C-terminal side (Asp effect) when singly charged ions are selected as a precursor ion[21]. These disadvantages of TMPP-ac are considered due to the absence of “mobile protons” in the derivatized peptide that can mediate the cleavage of amide bonds under low-energy CID conditions[22]. Thus, N-terminus-modifying groups with a certain level of affinity for protons are expected to be advantageous, in that they can be protonated during the ionization process, and that the protons added to the modified N-terminus can migrate to the amide bonds and promote their cleavage under low-energy CID conditions (Fig. 1). In this regard, a pyridyl group is one of the most suitable N-terminus modifiers for *de novo* peptide sequencing. The introduction of a pyridyl group, by virtue of its affinity for protons, enables the preferential detection of b-ions by retaining the proton in N-terminal fragments[15].

In spite of numerous attempts to find good derivatization tools for *de novo* peptide sequencing, the relationship between proton affinity and the fragmentation-promoting effect of a peptide/protein derivatization reagent under low-energy CID conditions has not been examined in detail. Also, the optimum physicochemical properties required for efficient fragmentation remain unclear. In this study, to explore the structure of charge derivatization appropriate for *de novo* sequencing, we prepared several reagents with

high proton affinity and examined their usefulness in *de novo* peptide sequencing under low-energy CID conditions.

Experimental section

Chemicals.

4-Amidinobenzamide hydrochloride and 6-amidino-2-naphthol methanesulfonate were purchased from Tokyo Chemical Industry (Tokyo, Japan). 4-(Aminomethyl)benzoic acid, 1*H*-pyrazole-1-carboxamidine hydrochloride, nicotinic acid and 4-methoxybenzylalcohol were obtained from Wako Pure Chemical Industries (Osaka, Japan). Methyl bromoacetate and *N*-hydroxysuccinimide were from Nacalai Tesque (Kyoto, Japan). 4-Nitrophenyl chloroformate was purchased from Aldrich (St. Louis, MO, USA). 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride was from Watanabe Chemical Industries (Hiroshima, Japan). 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM) was obtained from Kokusan Chemical (Tokyo, Japan). Angiotensin I was purchased from Calbiochem (Darmstadt, Germany). The other peptides used in this study were synthesized using Fmoc solid-phase chemistry in our laboratory. Nicotinoyloxysuccinimide was synthesized as described in the literature[15].

Synthesis of 4-amidinobenzoic acid.

4-Amidinobenzamide hydrochloride (5 g, 25 mmol) was dissolved in a mixture of 6 M HCl (150 ml) and 12 M acetic acid (30 ml) and the solution was refluxed at 110°C for 6 h.

The reaction solution was cooled with ice and the resulting precipitate was filtered to yield a crystal of 4-amidinobenzoic acid hydrochloride (4.6 g, 92%). ^{1}H-NMR (400MHz: DMSO-*d*6) 7.99 (d, 2H *J*=8.4 Hz), 8.15 (d, 2H *J*=8.2 Hz), 9.49 (s, 1H), 9.66 (s, 2H), 13.53 (br.s, 1H). ESI-MS: $[M+H]^{+}$ (*m/z*) calculated: 165.1, found: 165.1

Synthesis of 4-(guanidinomethyl)benzoic acid.

4-(Aminomethyl)benzoic acid (0.5 g, 3.3 mmol) was dissolved in 1 M sodium carbonate (10 ml). 1H-pyrazole-1-carboxamidine hydrochloride (0.5 mg, 3.4 mmol) was added, and the solution was stirred at room temperature overnight. Next, acetone (100 ml) was added, and the resultant precipitate was washed with acetone (x2), diethyl ether (x3) and DMF (x3). The precipitate was dissolved in a small amount of DMSO and reprecipitated by acetone after the removal of insoluble materials to yield 4-(guanidinomethyl)benzoic acid (0.2 g, 32%). ^{1}H-NMR (400MHz: DMSO-*d*6) 4.36 (d, 2H *J*=5.7 Hz), 7.22 (d, 2H *J*=8.2 Hz), 7.64 (br.s, 3H), 7.84 (d, 2H, *J*=8.2 Hz). ESI-MS: $[M+H]^{+}$ (*m/z*) calculated: 194.1, found: 194.1

Synthesis of 2-(6-amidino-2-naphthyloxy)acetic acid.

2-(6-amidino-2-naphthyloxy)acetic acid was synthesized as described previously.[23] Briefly, to a solution of 6-amidino-2-naphthol methanesulfonate (0.52 g, 1.8 mmol) and triethylamine (0.88 ml, 7 mmol) in DMF (10 ml) was added a solution of 4-methoxybenzyl-4-nitrophenylcarbonate (0.7 g, 2.3 mmol), which was prepared from the reaction of 4-methoxybenzylalcohol with 4-nitrophenyl chloroformate in the presence

of pyridine, in DMF (10 ml) dropwise over 5 minutes. The reaction mixture was stirred at 50°C for 3 h, and at room temperature for 15 h, poured into water (25 ml), acidified with cold 10% citric acid (4.5 ml) and extracted with dichloromethane (3 × 50 ml). The combined organic solution was washed with 10% sodium bicarbonate solution and brine, dried ($MgSO_4$), concentrated and purified by silica gel column chromatography (EtOAc/hexane, 1:3 to 1:1) to give *N*-(4-methoxybenzyloxycarbonyl)-6-amidino-2-naphthol (0.59 g, 94%). A solution of this compound (0.48 g, 1.4 mmol) in acetone (32 ml) was treated with potassium carbonate (0.64 g, 4.4 mmol) and methyl bromoacetate (0.38 g, 2.5 mmol), and stirred at room temperature for 48 h. The reaction mixture was concentrated, extracted with ethyl acetate, and washed with water and brine. The organic solution was dried, evaporated and purified by silica gel column chromatography (EtOAc/hexane, 1:1) to give methyl 4-[N-(4-methoxybenzyloxycarbonyl)-6-amidino-2-naphthyloxy]acetic acid (0.37 g, 63%). This compound (0.25 g, 0.6 mmol) was dissolved in a 1 M sodium hydroxide solution and the reaction mixture was stirred at room temperature for 1 h. The solution was acidified with 1 M HCl to pH 2 and filtered. The filtrate was evaporated, and the residue was dissolved in DCM (40 ml) and TFA (0.4 ml, 4 mmol). The reaction mixture was stirred at room temperature for 10 min. After the solvent was evaporated, diethyl ether was added, and the resultant solid was filtered, washed with diethyl ether and dried in vacuo to give 2-(6-amidino-2-naphthyloxy)acetic acid in quantitative yield. ^{1}H-NMR (400MHz, DMSO-*d*6) δ 4.86 (s, 2H), 7.39 (m, 2H), 7.83 (d, *J*=8.7 Hz, 1H) 7.99 (m, 2H)

8.51 (s, 1H) 9.40 (s, 2H) 9.53 (s, 2H) 13.14 (br.s, 1H). ESI-MS: $[M+H]^+$ (*m/z*) calculated: 245.1, found: 245.1

Synthesis of 4-amidinobenzoyloxysuccinimide.

4-Amidinobenzoic acid (328 mg, 2 mmol) was dissolved in dry DMF and mixed with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (383 mg, 2 mmol) and *N*-hydroxysuccinimide (230 mg, 2 mmol). The mixture was stirred under argon at room temperature overnight. The solution was evaporated *in vacuo*, and the residue was purified by silica gel column chromatography (EtOAc/$CHCl_3$/TFA, 80:20:0.5) to give 4-amidinobenzoyloxysuccinimide (170 mg, 33%). ^{1}H-NMR (400MHz, DMSO-*d*6) δ 2.92 (s, 4H), 8.03 (d, 2H, *J* = 8.2 Hz), 8.31(d, 2H, *J* = 8.3 Hz), 9.45 (br.s, 3H). ESI-MS: $[M+H]^+$ (*m/z*) calculated:262.1, found: 261.9

Peptide derivatization.

For derivatization with 4-amidinobenzoic acid or nicotinic acid, the succinimide ester of each compound was dissolved in phosphate buffer (100 mM, pH 7.5 for peptides without lysine residues and pH 5.9 for peptides with lysine residues) at a concentration of 10 mM, and mixed with peptides (1 mM) at room temperature overnight. The derivatized peptides were purified by reversed phase (RP) -HPLC (H_2O/CH_3CN with 0.1%TFA). For derivatization with 2-(6-amidinonaphthalen-2-yloxy)acetic acid or 4-(guanidinomethyl)benzoic acid, each compound was dissolved in DMSO at a concentration of 2.5 mM and mixed with peptides (0.25 mM) and DMT-MM (2.5 mM).

The resultant mixture was stirred at room temperature overnight. The derivatized peptides were purified by RP-HPLC (H_2O/CH_3CN with 0.1%TFA).

Mass spectrometric analysis.

The MS^n analysis was performed on a Shimadzu LCMS-IT-TOF mass spectrometer equipped with a nano-electrospray ion source (Kyoto, Japan). Samples were dissolved in 50% acetonitrile and 0.1% formic acid to a final concentration of ~10 μM and loaded into nanospray tips (HUMANIX, Hiroshima, Japan). All mass spectra were obtained in the positive mode, and the precursor ions for the analysis were manually selected. A potential of 1.0 kV was applied to the nanospray tip in the ion source. For CID experiments, argon was used as the collision gas and the collision energy was set at 120% (arbitrary units) with 30 ms of activation time both for MS^2 and MS^3. Precursor ion isolation width was set at 3.0 *m/z* units.

Computational calculations.

pK_b values were calculated using ACD/pKa DB (Advanced Chemistry Development, Toronto, Canada). Proton affinity was calculated using the following equation,

$$PA = -\Delta E_{elec} - \Delta ZPE + 5/2RT$$

where ΔE_{elec} and ΔZPE are the differences in electronic energies and zero-point energies, respectively, between protonated and unprotonated molecules. The Spartan'06 program

(Wavefunction, Irvine, CA) was used for these calculations. Geometry optimization was carried out at the HF/6-31G*, and energies were obtained from MP2/6-311G* single-point calculations. An amide form of each derivatization compound was used for the calculation (Table 2).

Results and discussion

Selection of derivatization groups with high proton affinity.

First we explored the structures having high proton affinity that could be suitable for N-terminal derivatization. The structural effect of N-terminal-modifying groups on the cleavage of peptide bonds in a molecule under low-energy CID conditions is likely implicated in its role as a "reservoir of mobile protons[22]". It has been proposed that the cleavage of a peptide bond under such conditions is driven by protonation at the nitrogen atom of amide bonds (Fig. 1), where the proton is transferred from N-terminal or other basic moieties in the molecule that was protonated by the initial ionization process[24]. After the bond cleavage, the proton at the oxazolone ring is again transferred back to an energetically more favorable site such as the N-terminal amino group, in order to form a b-ion. The proton during this process is referred to as a "mobile proton", and plays an important role in determining whether b- or y-ions are more favorably produced after the cleavage. If the proton affinity of the N-terminal modifying groups is high, the mobile proton is more readily transferred there to form b-ions.

So, the N-terminal-modifying groups should have high proton affinity, and the selection of a structure is made primarily based on the calculated pK_b values. Although pK_b is a

measure of basicity in an aqueous solution, it is highly correlated with the gas phase basicity, and therefore, suitable for the selection of compounds with high proton affinity. The availability and/or the ease of synthesis of derivatization reagents were also taken into consideration. Of many possible basic moieties, a guanidino moiety, a functional group in the Arg side chain, was chosen as a good candidate for the N-terminal charge derivatization, due to its low pK_b value (<1). An amidino moiety, which has a pK_b value slightly higher than that of a guanidino group (1~2), was also considered appropriate. The presence a carboxyl group in the derivatization reagent is essential for forming an amide linkage with the N-terminal amino group, and a benzene ring is incorporated to facilitate the synthesis of the reagents. Thus, we decided to test 4-guanidinomethylbenzoic acid (Gmb, pK_b=1.9) and 4-amidinobenzoic acid (Aba, pK_b=3.2) for derivatization (Fig. 2). For the latter, substitution of the benzene ring with a naphthalene ring was expected to further lower the pK_b value based on the calculation, and therefore, 2-(6-amidino-2-naphthyloxy)acetic acid (Ana, pK_b=2.6), which was relatively easily synthesized among the acidic compounds having a amidinonaphthyl moiety, was also tested in this study. Their performances as a derivatization reagent were evaluated based on a comparison with nicotinic acid (Nic, pK_b=9.2).

Comparison of effects on fragmentation among derivatives.

We first evaluated the effects of four derivatization reagents on fragmentation using model peptides (Table 1). Basic residues such as Arg, His and Lys often disrupt the fragmentation process by trapping mobile protons on their side chains. Therefore, we

evaluated the effect of derivatization using model peptides with or without basic amino acid residues. Peptides were derivatized at the N-terminal amino group by the reagents using DMT-MM as a condensation reagent[25], and analyzed by an ion trap time-of-flight hybrid mass spectrometer.

1. Peptide without basic amino acid residues.

Singly charged ions were chosen as precursor ions for CID experiments of peptide **1** (AAGLQIA), which contains no basic amino acid residues. When the peptide was analyzed without derivatization, a product ion spectrum shown in Fig. 3A was obtained. While several b- and y-ions were observed, not all of the diagnostic fragment ions required for determining the entire sequence of the peptide were obtained, even using MS^3. No b_1 ion was observed, as was expected given the proposed mechanism of peptide fragmentation in the gas phase, which requires a modification of the N-terminal amino group with an amide linkage. Derivatization of the N-terminus with nicotinic acid, which was reported to facilitate the identification of N-terminal fragments[15], clearly gave a whole series of b-ions including b_1 (Fig. 3B). However, some of the y-ions were still observed at an intensity significant to complicate the spectrum, and the relative intensity of b-ions in the low mass range (e.g., b_3 and b_4) was lower than that of the high mass range (e.g., b_6). When the peptide was derivatized with Gmb, Aba or Ana, fragmentation was improved in all cases, providing a whole series of b-ions (Fig. 3C-E). In contrast to the derivatives with nicotinic acid, the relative intensity of b_6 ions from all of these derivatives was

decreased, while that of b-ions in the low mass range was significantly increased. Unfortunately, neutral losses of ammonia from b-ions were markedly observed in the cases of derivatives with Gmb and Ana, which made the resultant product ion spectra of Gmb- and Ana-derivatized peptides relatively complicated. This hampered the identification of b-ions, and was considered unfavorable for the application of these derivatizations to other peptides. The loss of ammonia molecules is generally known to occur in the side chains of Arg, Lys, Gln and Asn. Since peptide **1** lacks these amino acid residues, the guanidino and amidino moiety of Gmb and Ana, respectively, were most likely the sources of ammonia. By contrast, the peaks arising from neutral losses of ammonia were not significantly observed in the product ion spectrum of the Aba-derivatized peptide, and this derivative was considered more advantageous in terms of the sequence analysis. It is known that the ammonia loss is driven by the protonation at nitrogen atoms in the guanidino moiety of Arg. It is likely that the relatively high proton affinity of the guanidino and amidino groups in Gmb and Ana, respectively, strongly stabilizes the protonated forms, increasing the chance of elimination of ammonia. The amidino group of Aba also has high proton affinity, but its strength does not seem to be high enough to eliminate ammonia molecules.

2. Peptide with basic amino acid residues.

The effect of derivatization was then evaluated using peptide **2** (DRVYIHPFH, angiotensin I), which contains three basic amino acid residues. Since singly charged ions were not generated from this peptide, doubly charged ions were chosen as a

precursor for the MS/MS analysis. Complete suppression of y-ions was difficult in this case, so the improving effect of derivatization was primarily assessed by the number of b-ions detected in the subsequent experiments.

In the product ion spectrum of the underivatized peptide (Fig. 4A), seven b-ions were observed. The amide bond between Ile and His was not cleaved by this fragmentation, which made *de novo* determination of the whole peptide sequence impossible. Derivatization with Nic did not increase the number of b-ions detected (Fig. 4B). In particular, b_1 and b_3 ions were not observed, indicating that the cleavage of amide bonds in the N-terminal region was not effective, although the detection of these ions might be possible by MS^3. Derivatization with Gmb and Ana also gave poor fragmentation, and the number of b-ions detected was not increased (Fig. 4D and 4E). In these cases, the generation of b_1 ions was most intense, whereas b_3, b_4 and b_5 ions were detected only weakly if at all. This indicates that the cleavage of amide bonds in the central region did not effectively occur in the Gmb- and Ana-derivatized peptides. By contrast, the introduction of Aba gave much more informative results (Fig. 4C), with whole series of b-ions observed in this spectrum, although the intensity of some of the y-ions (y_2 and y_4) was higher that that of b-ions. As a consequence, only the derivatization with Aba enabled the complete *de novo* sequencing of peptide **2**.

In the process of peptide fragmentation of singly charged ions, the mobile proton is usually supplied from the protonated N-terminal amino group. Thus, when the N-terminal amino group is modified by a charged group like a quarternary phosphonium,

the singly charged ion generated has no mobile protons, giving rise to poor fragmentation under low energy CID conditions. When a group with high proton affinity is introduced into the N-terminal amino group of a peptide, protonation during the ionization step occurs preferentially at the modified N-terminus, and this proton can serve as a mobile proton to mediate the cleavage of amide bonds. If the affinity is too high, the mobility of the proton is restricted to lower the efficiency in terms of the peptide bond cleavage. On the other hand, weak proton affinity is disadvantageous for the back transfer of protons to the modified N-terminus after bond cleavage, as well as for the initial protonation during the ionization step, resulting in less efficient fragmentation to form b-ions. Thus, an optimal proton affinity is required for N-terminal modifier groups to cause fragmentations adequate for the sequence analysis.

Given these considerations, unfavorable fragmentation of Gmb- or Ana-derivatized peptides is likely to be due to a rather high proton affinity of these moieties. Consequently, protons are localized to a greater extent at the N-terminus to cause a significant loss of ammonia from the derivatized N-terminus of peptide **1** and preferential fragmentations in the N-terminal region of peptide **2**. By contrast, the ineffective fragmentation at the N-terminal region of Nic-derivatized peptide **2** is possibly ascribable to the rather low affinity of the Nic moiety. In this case, the presence of substructures with high proton affinity in the C-terminal region such as a His side chain could have unfavorably affected the transfer of protons to the amide bonds in the N-terminal region. Thus, it is likely that Aba has the optimal proton affinity for peptide fragmentation under low-energy CID conditions.

The *ab initio* calculation of proton affinity supported these considerations (Table 2). In this study, considering a balance between accuracy and computational cost, we used the HF/6-31G* and MP2/6-311G* levels of theory for geometry optimization and single-point calculation, respectively. The calculated proton affinity of Nic in Table 2 is in good agreement with the experimental value[26], indicating that the calculation method is valid. As expected, all compounds have higher proton affinity than that of an amino group of alanine[27]. When comparisons were made between compounds, the proton affinity of Aba was intermediate, being higher than that of Nic but lower than that of Gmb or Ana. Such a "moderate" affinity is considered effective for the "catch" and "release" of protons during ionization and fragmentation (Fig. 1).

Establishment of a peptide derivatization method. Before the effect of Aba on the fragmentation of various peptides was evaluated, a facile and effective derivatization method was developed. For the formation of an amide linkage at the N-terminal amino group, a succinimide ester method has often been used[15]. Succinimide ester compounds are relatively stable and rapidly react with amino groups of peptides in an aqueous solution. Thus, the succinimide ester of Aba was synthesized, and reacted with the model peptide. As shown in Fig 5, almost complete derivatization of peptide **1** was achieved within 2 h when reaction was performed in a buffer at pH 7.5. All by-products of the reaction, such as 4-amidinobenzoic acid, are highly hydrophilic and so could be easily separated by RP-HPLC without interfering with the elution of derivatized peptides. While the succinimide ester of Aba can also react with an amino group of the Lys side

chain in the peptide in this reaction, specific introduction into the N-terminal amino group could be achieved by using a weakly acidic buffer (pH 5.9) for the reaction; however, complete modification at the N-terminal amino group was difficult under these conditions[15]. For selective modification at the N-terminal amino group, oxidative amide formation using alkynes could be a promising method[28].

Validation of the effect of 4-amidinobenzoic acid on fragmentation. The effect of Aba on fragmentation was further evaluated using four model peptides, **3**-**6**, which have 11-13 amino acid residues (Table 1).

1. Peptides without basic amino acid residues.

We first examined the effect of Aba-derivatization using two peptides containing no basic amino acid residues. A singly charged ion was selected as a precursor in each of the following analyses. Fragmentation of the underivatized peptide **3**, DYPVDIYYLMD, gave only seven b- and four y-ions (Fig. 6A), with intense fragmentation in the C-terminal region (e.g., b_9 and b_{10}). Since b_1 and b_3 ions were not observed, *de novo* sequencing of the peptide was impossible from this spectrum, even though the MS^3 analysis was performed. Peptide **4**, SYANGFSATSASL, also showed poor fragmentation without derivatization as shown in Fig. 6C (six b- and three y-ions). In addition, neutral losses of water and ammonia molecules from the produced fragment ions were frequently observed, as is often the case for peptides containing

Ser, Thr and Asn residues. The product ion spectrum eventually gave only limited information on the amino acid sequence of **4**.

In contrast, derivatization of these peptides with Aba, had a marked effect on the fragmentation, and a complete series of b-ions were successfully detected with the almost complete suppression of y-ion formation to enable the *de novo* sequencing of the peptides. The effect was particularly evident in the product ion spectrum of Aba-derivatized peptide **3** (Fig. 6B), in which all the b-ions were clearly detected with a similar level of intensity, except for the b_5 ion: the intensity of this ion was relatively high due to the Asp effect[29]. In the case of Aba-derivatized peptide **4** (Fig. 6D), neutral losses of water or ammonia molecules were still observed to some extent, but assigning the fragment ions was much easier than in the case of the underivatized peptide, which allowed complete sequencing.

2. Peptides with basic amino acid residues.

We next examined the effect of Aba-derivatization using peptides containing basic amino acid residues. A doubly charged ion was selected as a precursor in each of the following analyses. We first selected peptide **5** in which a Lys residue is located at the C-terminus, considering the use of this method for sequencing of tryptic peptides. Product ion spectrum of the underivatized peptide only showed the predominant bond cleavage between Phe and Pro residues (e.g., b_2 and y_4) because an amide bond N-terminal to a proline residue is particularly labile (Fig. 7A). However, when the peptide was derivatized with Aba, a complete series of b-ions were detected in the

product ion spectrum (Fig. 7B), which allowing easy interpretation of the spectrum. Although an amino moiety of the Lys side chain can also react with the succinimide ester of Aba, the specific introduction of 4-amidinobenzoic acid into the N-terminal was achieved by adjusting the reaction buffer to pH 5.9. We then performed the MS/MS analysis of peptide **6**, ATQQTAAYKTLVS, without derivatization. Nine b- and six y-ions were observed (Fig. 7C) to preclude determination of the sequence in the N-terminal region. This was likely because of the high basicity of the Lys residue in the C-terminal region. After derivatization with Aba, a complete series of b-ions were clearly detected in the product ion spectrum (Fig. 7D), although MS^3 analysis was required to detect the b_1 ion. Identification of some of the b-ions was difficult due to the overlap with y- and other ions in the spectra, but they were clearly distinguished by comparison of spectra between underivatized and derivatized peptides, because *m/z* values of y-ions remain unchanged after derivatization.

A complete series of b-ions were also detected for peptide **7**, TDVNGDGRHAL, after derivatization with Aba (Fig. 7F), whereas only incomplete fragmentation occurred before derivatization (Fig. 7E). In the product ion spectrum of the Aba-derivatized peptide, the intensity of b_7 and b_8 ions was still low. This is because of the high proton affinity of the Arg side chain, which is known to suppress the dissociation of amide bonds adjacent to the Arg residue. It has been shown that such an unfavorable effect of the Arg residue can be removed by modification of a guanidino group of the Arg side chain with acetylacetone[30] or malondialdehyde[31].

Conclusions

In this study, we evaluated the improving effect of N-terminal charge derivatization on peptide fragmentation using compounds with high-proton affinity. Comparison of the effects on fragmentation among four derivatization reagents differing in proton affinity clearly indicated that there was an optimal affinity for efficient fragmentation of peptides: a balance between the release and catch of protons for the activation of amide bonds and for the preferential generation of b-ions, respectively, is needed to achieve complete fragmentation. Among four reagents tested, derivatization with 4-amidinobenzoic acid brought about the most effective fragmentation in all peptides used in this study, which enabled *de novo* peptide sequencing; however, the identification of b-ions was still not very easy in some cases due to the overlap of b-ions with other series of ions, especially when multiply charged ions were selected as a precursor. The utilization of isotope-labeled derivatization reagents may provide a much easier solution for the identification of b-ions by comparing the spectra between the derivatives obtained by labeled and non-labeled reagents[15]. Experiments using ^{15}N-labeled 4-amidinobenzoic acid are currently underway. Derivatization with 4-amidinobenzoic acid will undoubtedly improve the peptide fragmentation pattern, which may offer a novel *de novo* peptide sequencing method under low-energy CID conditions.

Acknowledgment

This study was supported in part by 21st century COE program for Innovative Food and Environmental Studies Pioneered by Entomomimetic Sciences of The Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

[1] D.F. Hunt, J.R. Yates, 3rd, J. Shabanowitz, S. Winston, C.R. Hauer. Protein sequencing by tandem mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233.

[2] M. Wilm, A. Shevchenko, T. Houthaeve, S. Breit, L. Schweigerer, T. Fotsis, M. Mann. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry, *Nature*. **1996**, *379*, 466.

[3] V.H. Wysocki, K.A. Resing, Q. Zhang, G. Cheng. Mass spectrometry of peptides and proteins, *Methods*. **2005**, *35*, 211.

[4] R.S. Johnson, M.T. Davis, J.A. Taylor, S.D. Patterson. Informatics for protein identification by mass spectrometry, *Methods*. **2005**, *35*, 223.

[5] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell. Probability-based protein identification by searching sequence databases using mass spectrometry data, *Electrophoresis*. **1999**, *20*, 3551.

[6] J.K. Eng, A.L. Mccormack, J.R. Yates. An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database, *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976.

[7] L.Y. Geer, S.P. Markey, J.A. Kowalak, L. Wagner, M. Xu, D.M. Maynard, X.Y. Yang, W.Y. Shi, S.H. Bryant. Open mass spectrometry search algorithm, *J. Proteome Res.* **2004**, *3*, 958.

[8] J. Seidler, N. Zinn, M.E. Boehm, W.D. Lehmann. De novo sequencing of peptides by MS/MS, *Proteomics*. **2010**, *10*, 634.

[9] L.D. Fricker, J. Lim, H. Pan, F.Y. Che. Peptidomics: identification and quantification of endogenous peptides in neuroendocrine tissues, *Mass Spectrom. Rev.* **2006**, *25*, 327.

[10] H.J. Cooper, K. Hakansson, A.G. Marshall. The role of electron capture dissociation in biomolecular analysis, *Mass Spectrom. Rev.* **2005**, *24*, 201.

[11] L.M. Mikesh, B. Ueberheide, A. Chi, J.J. Coon, J.E. Syka, J. Shabanowitz, D.F. Hunt. The utility of ETD mass spectrometry in proteomic analysis, *Biochim. Biophys. Acta*. **2006**, *1764*, 1811.

[12] S.E. Martin, J. Shabanowitz, D.F. Hunt, J.A. Marto. Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC micro-ESI fourier transform ion cyclotron resonance mass spectrometry, *Anal. Chem.* **2000**, *72*, 4266.

[13] A.M. Frank, M.M. Savitski, M.L. Nielsen, R.A. Zubarev, P.A. Pevzner. De novo peptide sequencing and identification with precision mass spectrometry, *J. Proteome Res.* **2007**, *6*, 114.

[14] K.D. Roth, Z.H. Huang, N. Sadagopan, J.T. Watson. Charge derivatization of peptides for analysis by mass spectrometry, *Mass Spectrom. Rev.* **1998**, *17*, 255.

[15] M. Munchbach, M. Quadroni, G. Miotto, P. James. Quantitation and facilitated de novo sequencing of proteins by isotopic N-terminal labeling of peptides with a fragmentation directing moiety, *Anal. Chem.* **2000**, *72*, 4047.

[16] P.L. Ross, Y.N. Huang, J.N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D.J. Pappin. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents, *Mol. Cell. Proteomics*. **2004**, *3*, 1154.

[17] B. Spengler, F. Luetzenkirchen, S. Metzger, P. Chaurand, R. Kaufmann, W. Jeffery, M. Bartlet-Jones, D.J.C. Pappin. Peptide sequencing of charged derivatives by postsource decay MALDI mass spectrometry, *Int. J. Mass Spectrom.* **1997**, *169*, 127.

[18] D.S. Wagner, A. Salari, D.A. Gage, J. Leykam, J. Fetter, R. Hollingsworth, J.T. Watson. Derivatization of peptides to enhance ionization efficiency and control fragmentation during analysis by fast atom bombardment tandem mass spectrometry, *Biol. Mass Spectrom.* **1991**, *20*, 419.

[19] L.N. Marekov, P.M. Steinert. Charge derivatization by 4-sulfophenyl isothiocyanate enhances peptide sequencing by post-source decay matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, *J. Mass Spectrom.* **2003**, *38*, 373.

[20] Z.H. Huang, J. Wu, K.D. Roth, Y. Yang, D.A. Gage, J.T. Watson. A picomole-scale method for charge derivatization of peptides for sequence analysis by mass spectrometry, *Anal. Chem.* **1997**, *69*, 137.

[21] C. Gu, G. Tsaprailis, L. Breci, V.H. Wysocki. Selective gas-phase cleavage at the peptide bond C-terminal to aspartic acid in fixed-charge derivatives of Asp-containing peptides, *Anal. Chem.* **2000**, *72*, 5804.

[22] V.H. Wysocki, G. Tsaprailis, L.L. Smith, L.A. Breci. Mobile and localized protons: a framework for understanding peptide dissociation, *J. Mass Spectrom.* **2000**, *35*, 1399.

[23] C. Bailey, E. Baker, J. Hayler, P. Kane. Amidine protection for solution phase library synthesis, *Tetrahedron Lett.* **1999**, *40*, 4847.

[24] N.C. Polfer, J. Oomens, S. Suhai, B. Paizs. Infrared spectroscopy and theoretical studies on gas-phase protonated leu-enkephalin and its fragments: direct experimental evidence for the mobile proton, *J. Am. Chem. Soc.* **2007**, *129*, 5887.

[25] M. Kunishima, A. Kitao, C. Kawachi, Y. Watanabe, S. Iguchi, K. Hioki, S. Tani. A racemization test in peptide synthesis using 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM), *Chem. Pharm. Bull.* **2002**, *50*, 549.

[26] E.P.L. Hunter, S.G. Lias. Evaluated gas phase basicities and proton affinities of molecules: An update, *J. Phys. Chem. Ref. Data*. **1998**, *27*, 413.

[27] A.G. Harrison. The gas-phase basicities and proton affinities of amino acids and peptides, *Mass Spectrom. Rev.* **1997**, *16*, 201.

[28] W.K. Chan, C.M. Ho, M.K. Wong, C.M. Che. Oxidative amide synthesis and N-terminal alpha-amino group ligation of peptides in aqueous medium, *J. Am. Chem. Soc.* **2006**, *128*, 14796.

[29] G. Tsaprailis, H. Nair, A. Somogyi, V.H. Wysocki, W.Q. Zhong, J.H. Futrell, S.G. Summerfield, S.J. Gaskell. Influence of secondary structure on the fragmentation of protonated peptides, *J. Am. Chem. Soc.* **1999**, *121*, 5142.

[30] S. Dikler, J.W. Kelly, D.H. Russell. Improving mass spectrometric sequencing of arginine-containing peptides by derivatization with acetylacetone, *J. Mass Spectrom.* **1997**, *32*, 1337.

[31] A. Foettinger, A. Leitner, W. Lindner. Derivatisation of arginine residues with malondialdehyde for the analysis of peptides and protein digests by LC-ESI-MS/MS, *J. Mass Spectrom.* **2006**, *41*, 623.

[32] A.G. Harrison. The gas-phase basicities and proton affinities of amino acids and peptides, *Mass Spectrom. Re.v.* **1997**, *16*, 201.

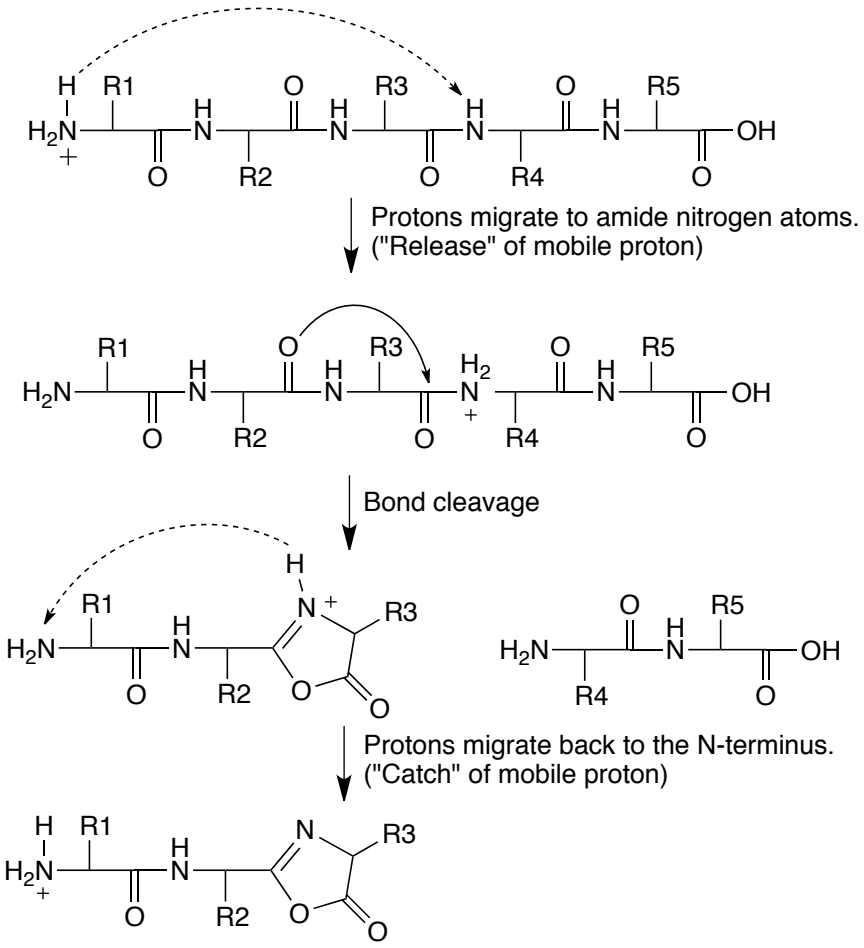

Fig. 1 Mechanism of formation of b-ions during peptide fragmentation

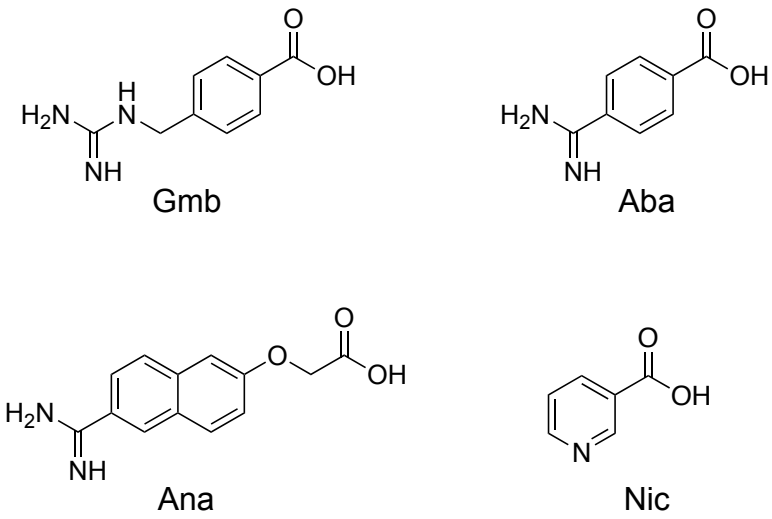

Fig. 2 Structure of derivatization reagents used in this study

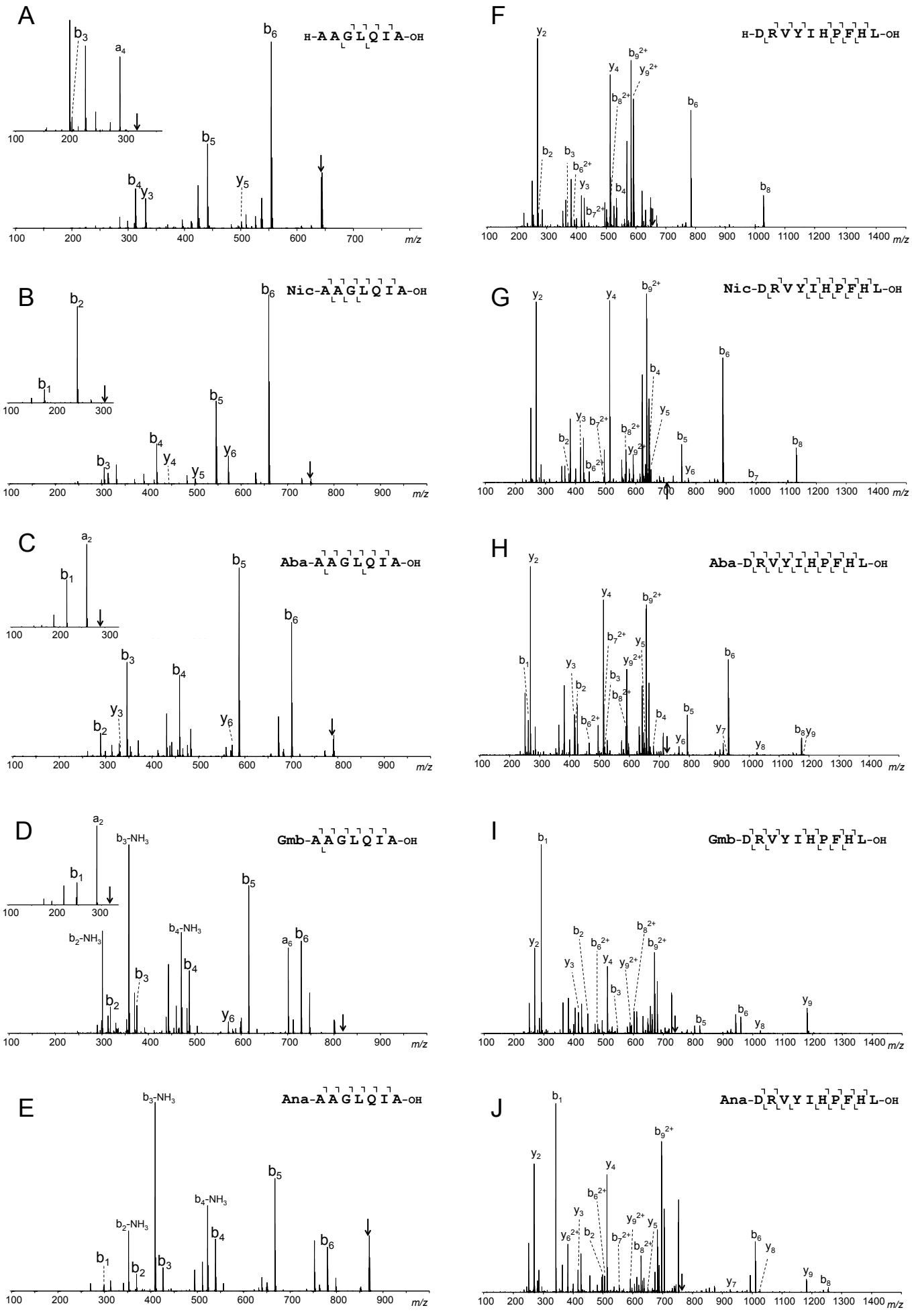

Fig. 3 Comparison of improving effect of N-terminal derivatization reagents on fragmentation. Arrows indicate precursor ions.

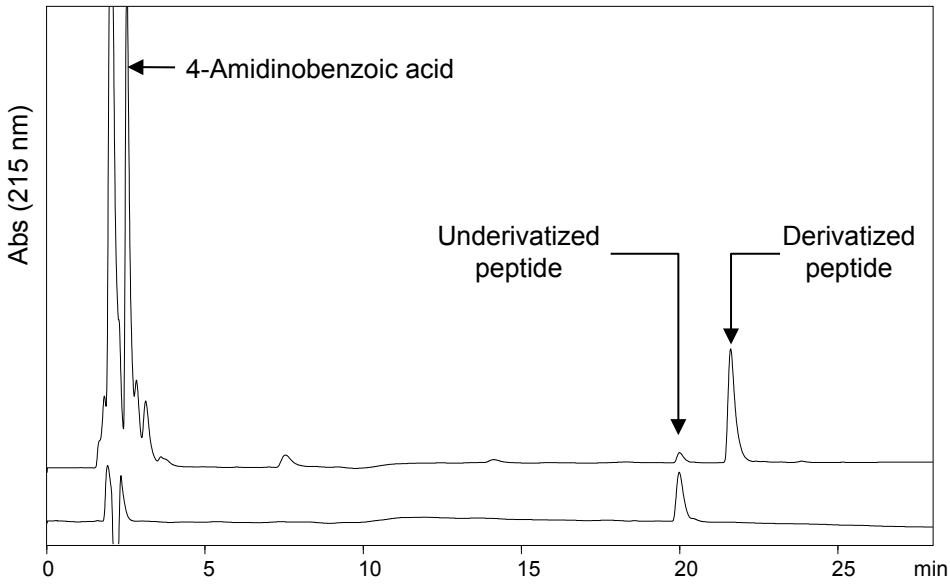

Fig. 4 HPLC analysis of the N-terminal derivatization reaction using succinimide ester of 4-amidinobenzoic acid. (Upper trace: reaction mixture after 2 h, lower trace: underivatized peptide)

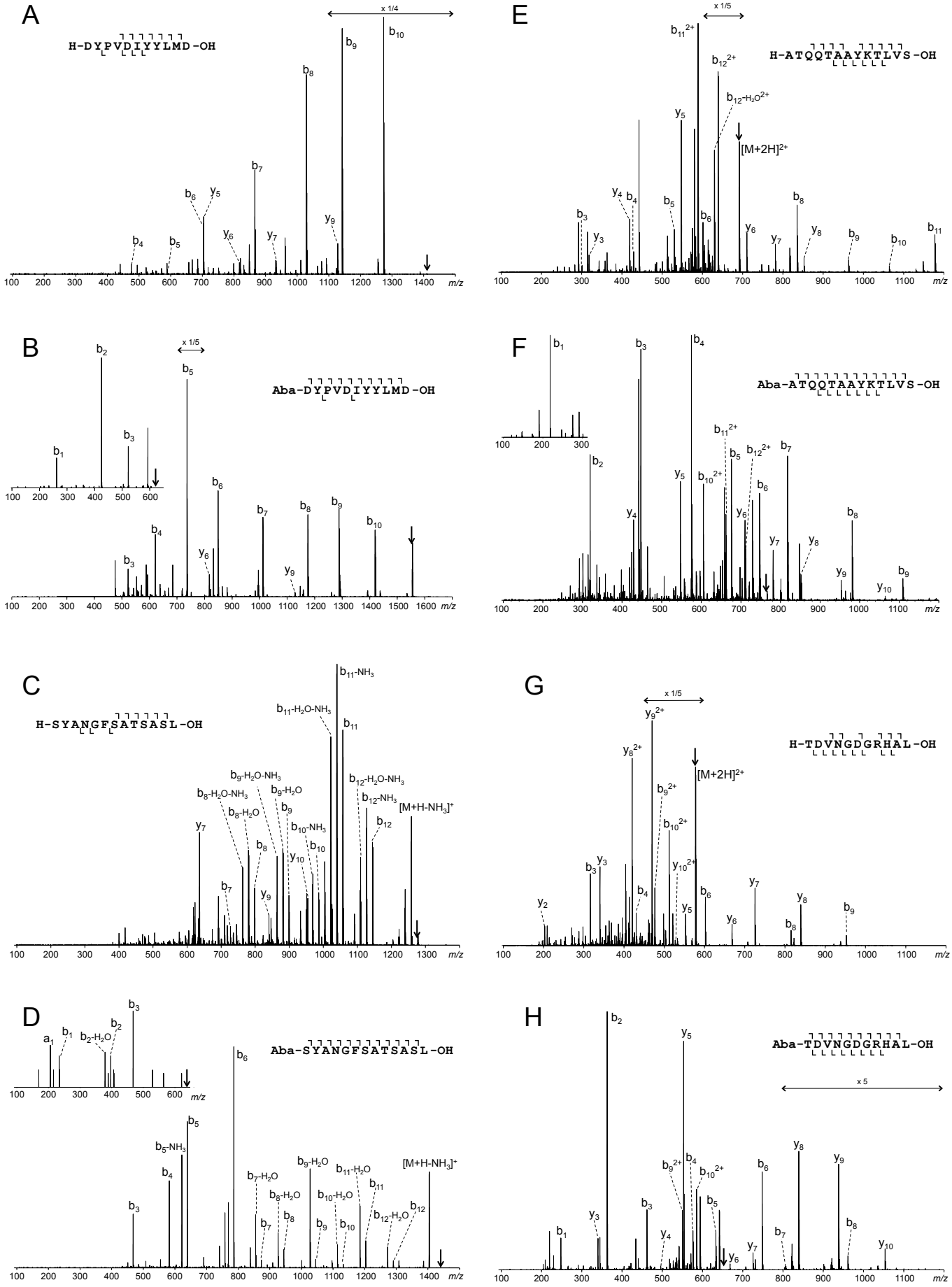

Fig. 5 Improving effect of 4-amidinobenzoic acid on fragmentation of various peptides

Table 1 Amino acid sequence of peptides used in this study

No.	Peptide sequence
1	AAGLQIA
2	DRVYIHPFHL
3	DYPVDIYYLMD
4	SYANGFSATSASL
5	ATQQTAAYKTLVS
6	TDVNGDGRHAL

Table 2 Calculated proton affinities of compounds used for derivatization

Compounds	PA (kcal/mol)
Gmb	244.5
Ana	241.4
Aba	235.2
Nic	215.3